

Creation of a global human rights disabilities media watch environment using a grid computing infrastructure

Jeffrey L. Tilson^a, Ezra Zubrow^b, Marcia H. Rioux^d, Raymond D. Cohen^e, James O. Whitlock^c

^aCenter for Computational Research

^bDepartment of Anthropology

^cComputer and Information Technology

University at Buffalo

{jtilson, zubrow, [whitlock](mailto:whitlock@buffalo.edu)}@buffalo.edu

^dSchool of Health Policy and Management

York University

mrrioux@yorku.ca

^eCanadian Abilities Foundation

Toronto, Ca

ray@abilities.ca

Abstract

The proposed work begins the specification of an e-Science[1] in direct support of the larger initiative, *Disability Rights Promotion International*, to establish a comprehensive international monitoring system for disability rights. Collaborating with disability rights organizations, human rights advocates, and the United Nations (UN) system. Adoption of modern hardware and, particularly, software advances are exploited to facilitate a large multinational distributed query and storage environment. Adoption of the Globus Toolkit® and use of a grid-services model simplify development and deployment.

Introduction

The United Nations estimates that 600 million people (10% of the world's population) have some form of disability[2] that potentially results in barriers to full participation in society and placing such people at greater risk of discrimination, abuse, and poverty. Governments' and Agencies' difficulties in recognizing both the needs for, and barriers to, equal rights serve to perpetuate the social isolation and exclusion often experienced by people with disabilities.

The United Nations Commission on Human Rights recently passed resolutions for the protection of the rights of people with disabilities and, as a result, adopted resolution 1998/31[3]. The resolution makes a series of statements and recommendations for future development in this area. The adoption of a subsequent resolution[4] reaffirms Resolution 1998/31 and calls for measures to strengthen the protection and monitoring of the human

rights of persons with disabilities. In December 2001, resolution 56/168 was adopted, establishing an Ad Hoc Committee to consider proposals for an international convention to protect and promote the rights and dignity of people with disabilities. Subsequent Ad Hoc committee meetings produced a draft Convention in January 2004.

Concurrently with the process to draft a disability rights convention, opportunities exist to develop a disability dimension in existing international, regional and national human rights systems. At a meeting hosted by the UN High Commissioner for Human Rights, the participants agreed on a multifaceted approach to disability.[5] They endorsed a "twin-track approach" whereby the drafting of a new convention would be seen as complimentary to strengthening attention to disability within the existing international human rights system".

The Disability Rights Promotion International.[6] (DRPI) organization emerged from a seminar[7] on human rights and disability held at the Almåsa Conference Centre in Stockholm. [8] Experts from all world regions examined measures to strengthen the protection and monitoring of the human rights of people with disabilities. The resolution makes a series of statements and recommendations that include a need, interpretable as a cyberinfrastructure [9], to support the monitoring of a broad range of media. Computational grid technology has now matured to the point where development and deployment of a disciplinary-focused grid infrastructure to a new and historically non-

technical field can be credibly conceived as a pathway to solution.

System Requirements

The use-case of the required environment is fairly simple in scope. It has been determined that the system needs to enable the efficient finding, acquiring, and storing of reports, news (video/audio/photos/text), and law-cases pertaining to disabilities issues worldwide. Support of independent correspondence (off-the-street reports) will be another necessity. These data will be archived for long periods of time and must be accessible in such a way that machine processing can be performed. The data must maintain its integrity and have a well defined chain-of-custody. The data are expected to be processed (mined) in obvious ways, but also to be “fused” as well. The fusion process[10] will provide necessary knowledge on information relevance, situational awareness, intent assessment, and ultimately augmenting the available information with simulation. Fusion of the data is an important capability as disability related data are inherently geographical.[11] Enabling these capabilities have been deemed necessary in the areas of Policy, Research, and Education. The typical access models estimated for each area are collected in Table 1.

Policy making is expected to be initially from within the health domain but other governmental agencies can conceivably benefit from such capability. The Canadian Gov’n has already demonstrated its commitment to this process by explicitly addressing disabilities in the development of their new health policy[12]. Occasional high-presence collaboration is expected with, typically, asynchronous (fusion) computation on the stored media data required.

In the *academia* area, a lecture format (talking-head) interaction must be supported with primarily read-only access to the stored media data. The archived data access are expected to be, typically, synchronously accessed.

The *research* area will be the most active work group within the system. These members indicate a pressing need to interact collaboratively across broad geographic

distances on a frequent basis. They also require access to the system’s data within their collaborative sessions. The media data access is both synchronous and asynchronous. This group develops tools (services) for managing and analyzing the data. Thus, we need to support shared data and instantiation of shared services within the collaboration framework.

Fraction of use (%)	Syn-chronous	Asyn-chronous	Real-time
Policy	20	60	20
Research	25	25	50
Education	60	30	10

Table 1. Approximate distribution of data access types for the three principle areas. Policy users are expected to spend the majority of the time fusing data (computationally intensive). Education users will be performing online data mining and occasional online lectures. The research area has the greatest needs with a substantial fraction of high-presence/immersive collaboration coupled with online data-mining and fusion.

As the relative complexities of the data access models vary, the system needs to support the extremes of each kind. Thus, substantial amounts of computing are required for data fusion, A high-performance procedure for accessing data online to support the synchronous access and the ability to support (real-time) high-presence collaborative sessions containing several video/audio streams are necessary. These requirements are predicted to grow substantially in the first few years.

Data type (stored)	Name	(typ) Speed (Mbps)
Video	SDI	1500
	DV (ave)	20-25
	DVD	6.0 (nom)
Audio (mono)	48 Khz/16 b	0.7
(stereo)	48 Khz/16 b (stereo)	1.4

Table 2. A representation of typical video/audio characteristics. SDI is illustrative of very high quality (production level) video and is for comparison only.

The ability to collect, format, and assimilate the desired kinds media information is hampered by several inadequacies. First is the lack of a coordinated technology effort to architect and deploy a secure and distributed data-collection framework for collecting and safely storing the relevant media (data) in a way amenable to search and subsequent analysis. The second is the standardization and deployment of data-templates that serve to define relevant media and interoperate with software data collection agents to select, format, and interact with the grid-framework and to assimilate the collected data back into the system. To deploy a prototype infrastructure sufficiently capable for preliminary testing, benchmarking, optimization, and real-life data collection suitable for testing with the DRPI, requires making progress on both of the stated inadequacies.

The first technical stage of the proposed work is to develop a technology solution to enable the monitoring of broad "main stream" media reporting on human right abuses of people with disabilities. The second stage is to begin the actual media monitoring (watching) and deploy research tools to analyze the collected data. The third stage is a pervasive integration of the developed tools into curricula at the participating institutions.

Infrastructure description

The needs, as specified by the experts, can be addressed using a computational grid infrastructure[13,14,15,16]. The Open Grid Services Architecture (OSGA) [17] and the Web Service Resource Framework (WSRF)[18] serve to define the grid services[19] environment within which this system is being developed. The Globus Toolkit@[20,21] is being exploited to implement this grid. We are resorting to our own virtual organizational (VO) [22] structure the access details of which will ultimately be defined by the DRPI. Several technical solutions exist that, in principle, can be assembled into a system that seeds the solution for this problem. The solutions selected for beginning the construction of the disabilities-grid begin with widely accepted technologies

enabling collaborative environments via the Access Grid™.[23,24,25]

The system is a basic client-server architecture centered about the UB Center for Computational Research and York University's Centre for Health Research. These sites will collectively act as a storage system for the disability-grid. In addition, they serve as a download site for new software such as new data-collection services. These geographically distant server sites will act as a data mirrors potentially, via a wide area GPFS filesystem that supports instantiated services running on servers at either site. The basic nature of the clients is such that we require several (internationally) distributed client (data-aggregator) sites each with their own locally described policies, participating in the generation of the large shared pool of information. Recent definitions used to define a grid [26] indicate that indeed, the human-rights infrastructure is a grid. Moreover, as recently quoted by Ian Foster[27] "...Ultimately the Grid must be evaluated in terms of the applications, business value, and scientific results that it delivers, not its architecture."

Resource Requirements

The network-centric nature of the system requires some consideration of issues in computation, storage, displays, and networking.

Computation

The level of computation (mining/fusing) will grow substantially over the lifetime of the project. Deployment is simplified, however, by the needs of the computational algorithms. Namely, massive parallelism is not required. It is only necessary to support many concurrent computations. This can be satisfied by resorting to racks of commodity servers accessible to the grid factory service. Compute nodes will be available at both UB and YorkU sites with load balancing of the tasks. This is facilitated by the unified file system across sites.

Storage

Substantial amounts of storage will ultimately be required over the lifetime of the system. This includes

both raw collected data and post-processed information consisting of mined, fused, and simulated results. We estimate a yearly store of $O(10)$ TB per year of raw data. Synthesized results will add to this. Access to the database itself must be highly reliable and so a two-site storage system has been selected. Testing of specific inter-site networking needs to be performed, but will support redundant-write I/O operations at each site without significantly impairing performance. Initially, a high performance shared network will be utilized. In the out-years, however, a (middleware) provision-able lambda-based network is anticipated.

Display

The participant sites require the ability to interact via large-scale collaboration. The digital-divide hampers some of the client sites access but cannot be a total impediment to connection. Display technology at the better equipped sites will need to support high resolution video/photos. Multiple displays will be required to satisfy the high-presence requirement. Multicasting must be an option. We expect as many as $O(30)$ simultaneous streams in addition to playing videos, etc.

So at least 1 Gbps networking would be beneficial. As many of the world's experts in disabilities require alternative modalities to participation, it is imperative to have basic translators (e.g., text-speech, etc.) available.

Networking

The networking needs are somewhat varied. The server and primary education sites require high-capacity connectivity to enable the full fidelity of the system. This includes the two-site (storage) filesystem, and the serving of data from the databases to the compute nodes in a synchronous manner. At least 1 Gbps is anticipated except for the inter-site link, which may need to be greater. The connectivity at the client sites is somewhat problematic as we have little control over this. Some sites may only have a POTS as a result of the digital-divide. While this doesn't substantially impact asynchronous analysis, it will impact the collaboration-technology itself. We must deploy a system that can function at some minimal level at such low capability.

The possibility of transcoding of the data/session is under exploration as a possible solution for these cases.

Conclusion

The disabilities-grid environment is in its formative stages but addresses a well-defined and widely acknowledge need stemming from the UN. A possible solution based on WS-RF is being prepared using the Globus Toolkit® and exploits techniques demonstrated to be successful for many scientific endeavors.

References

- [1] e-Science as used in this proposal is defined in the final report to the NSF, NSF CISE Grand Challenges in e-Science Workshop Report April, 2002. p 44.
(<http://www.evl.uic.edu/activity/NSF/index.html>. Namely, "...e-Science refers to large-scale science carried out through distributed global collaborations enabled by networks...")
- [2] Gerard Quinn and Theresia Degener, Human Rights and Disability: The Current Use and Future Potential of United Nations Human Rights Instruments in the Context of Disability (Geneva: Office of the High Commissioner for Human Rights, 2002) at page 1. See also the "Human Rights and Disability" web pages on the Office of the High Commissioner for Human Rights web site at <http://193.194.138.190/disability/index.htm>.
- [3] UN CHR res. 1998/31, ESCOR Supp. (No. 3) at 117, U.N. Doc. E/CN.4/1998/31 (1998).
- [4] UN CHR res. 2000/51, UN Doc. E/CN.4/Res/2000/51 (2000).
- [5] Report of the United Nations High Commissioner for Human Rights and Follow-up to the World Conference on Human Rights: Human rights of persons with disabilities E/CN.4/2002/18/Add.1, 12 February 2002.
<http://www.yorku.ca/drpi>.
- [7] This conference was convened by one of the project Co-PIs, Dr. Bengt Lindqvist, UN Special Rapporteur on Disability.
- [8] See Let the World Know: Report of a Seminar on Human Rights and Disability at
<http://www.un.org/esa/socdev/enable/stockholmnov2000.htm>

[9] Revolutionizing Science and Engineering Through Cyberinfrastructure: Report on the National Science Foundation Blue-Ribbon Panel on Cyberinfrastructure. Daniel E. Atkins, Chair, January 2003.

[10] Adapted from information by J. Llinas, Center for Multisource Information Fusion, The University at Buffalo. (<http://www.infofusion.buffalo.edu>)

[11] M. Rioux, E. Zubrow, M.S. Bunch, W. Miller, "Atlas of Literacy and Disability", Canadian Abilities Foundation, 2003.

[12] E. Zubrow, M. Rioux. Private communication, 2004.

[13] I. Foster and C. Kesselman, "The Grid: Blueprint for a New Computing Infrastructure," Morgan Kaufman, 1999.

[14] I. Foster, "The Grid: A New Infrastructure for 21st Century Science," *Physics Today*, vol. 55, pp. 42-47, 2002.

[15] I. Foster, "The Grid: Computing without Bounds" *Scientific American*, vol. 288(4), pp. 78-85, 2003.

[16] F. Berman, G. Fox, and T. Hey, *Grid Computing: Making Global Infrastructure a Reality*, John Wiley & Sons Ltd, West Sussex, England, (2003)

[17] I. Foster, C. Kesselman, J. Nick, S. Tuecke, "The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration." Open Grid Service Infrastructure WG, Global Grid Forum, June 22, 2002. <http://www.globus.org/research/papers/ogsa.pdf>

[18] Modeling Stateful Resource with Web Services, I. Foster and S. Tuecke, Eds. Version 1.1, 2004.

[19] S. Tuecke, K. Czajkowski, I. Foster, J. Frey, S. Graham, C. Kesselman, "Grid Service Specification," Open Grid Service Infrastructure Working Group, Global Grid Forum, 2002.

[20] I. Foster and C. Kesselman, "Globus: A Metacomputing Infrastructure Toolkit," *Intl J. Supercomputer Applications*, vol. 11, pp. 115-128, 1997.

[21] I. Foster and C. Kesselman, "The Globus Project: A Status Report," *IPPS/SPDP '98 Heterogeneous Computing Workshop*, 1998, pp. 4-18.

[22] I. Foster, C. Kesselman, S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations," *International Journal Supercomputing Applications*, 15, 2001.

[23] AccessGrid. <http://www.accessgrid.org>

[24] L. Childers, T.L. Disz, M. Hereld, R. Hudson, I. Judson, R. Olsen, M. E. Papka, J. Paris, R. Stevens, "Active Spaces on the Grid: The Construction of Advanced Visualization and Interactive Environments," *Paralleldatorcentrum Kungl Tekniska Hogskolan Seventh Annual Conference (Simulation and Visualization on the Grid)* vol. 13, Lecture Notes In Computational Science and Engineering, B. Engquist, L. Johnsson, M. Hammill, and F. Short, Eds. Stockholm, Sweden: Springer-Verlag, 1999, pp. 64-80.

[25] T.L. Disz, R. Evard, M.W. Henderson, W. Nickless, R. Olsen, M.E. Papka, R. Stevens, "Designing the Future of Collaborative Science: Argonne's Futures Laboratory," *IEEE Parallel and Distributed Technology Systems and Applications*, vol. 3, pp. 14-21, 1995.

[26] I. Foster "What is the Grid? A Three Point Checklist," *GRIDToday*, July 20, 2002.

[27] I. Foster, "Going for the Grid," *Computer Business Review Online*, 2004. http://www.cbronline.com/content/COMP/magazine/Articles/Servers_Mainframes/000223.asp